FEATURE LEVEL SENSOR FUSION

Tamar Peli, Mon Young, Robert Knox, Ken Ellis, Fredrick Bennett
Atlantic Aerospace Electronics Corporation
470 Totten Pond Road, Waltham, MA 02154

ABSTRACT

This paper describes two practical fusion techniques (hybrid fusion and cued fusion) for automatic target cueing that combine features derived from each sensor data at the object-level. In the hybrid fusion method each of the input sensor data is prescreened (i.e. Automatic Target Cueing (ATC) is performed) before the fusion stage. The cued fusion method assumes that one of the sensors is designated as a primary sensor, and thus ATC is only applied to its input data. If one of the sensors exhibits a higher Pd and/or a lower false alarm rate, it can be selected as the primary sensor. However, if the ground coverage can be segmented to regions in which one of the sensors is known to exhibit better performance, then the cued fusion can be applied locally/adaptively by switching the choice of a primary sensor. Otherwise, the cued fusion is applied both ways (each sensor as primary) and the outputs of each cued mode are combined. Both fusion approaches use a back-end discrimination stage that is applied to a combined feature vector to reduce false alarms. The two fusion processes were applied to spectral and radar sensor data and were shown to provide substantial false alarm reduction. The approaches are easily extendable to more than two sensors.

1.0 INTRODUCTION

Recent crisis and conflict operations have reinforced the need for broad area imagery coverage to support all stages of operations. The first utility of imagery usually takes the form of target detection/recognition and change detection. The ability to reliably and rapidly detect, discriminate and classify military targets can provide a significant tactical advantage in the battlefield. Automatic target detection and recognition (ATD/ATR) has been a focus of research for the last two decades. The performance of automatic target recognition has not yet reached the required level of recognition accuracy and speed. The complexity of the recognition process has forced the development of automatic target prescreening technologies in order to cue the complex/time-consuming recognition stage to limited/reduced data. Currently ATR processes are primarily used as a second layer for reduction of false cues (i.e., target / no target decision) and leaving the actual recognition to human operator.

The limitations (high probability of detection/recognition at an unacceptable level of false alarms) of current systems utilizing a single sensing domain in addressing the various deployed CC&D techniques have led to the incorporation of multiple sensors. It is expected that the result of fusing data from multiple independent sensors will offer the potential for better performance than can be achieved by either sensor, and will reduce vulnerability to sensor-specific countermeasures and deployment factors. The first and most significant increment in performance improvement will come from multi-source fusion at the prescreening (target detection) stage. This will enable either the use of current recognition processing or will allow the application of more complex processes to a reduced set of detections.

Multi/hyper-spectral (MS/HS) sensors are examples of the trend to add dimensionality (in this case - spectral) in order to achieve significant performance gains. This trend is reflected in plans to host multiple pods such as SAR, EO/IR, and multi-spectral sensors on UAVs. Multi/hyper-spectral and SAR sensor data offers the most potential for defeating CC&D. SAR sensors have the advantage of excellent standoff, all weather capabilities, and broad area

Form SF298 Citation Data

Report Date *("DD MON YYYY")* 00001999	Report Type N/A	Dates Covered (from... to) *("DD MON YYYY")*

Title and Subtitle Feature Level Sensor Fusion	**Contract or Grant Number**
	Program Element Number
Authors Tamar, Peli; Young, Mon; Knox, Robert; Ellis, Ken; Bennett, Fredrick	**Project Number**
	Task Number
	Work Unit Number
Performing Organization Name(s) and Address(es) Atlantic Aerospace Electronics Corporation 470 Totten Pond Road Waltham, MA 02154	**Performing Organization Number(s)**
Sponsoring/Monitoring Agency Name(s) and Address(es)	**Monitoring Agency Acronym**
	Monitoring Agency Report Number(s)
Distribution/Availability Statement Approved for public release, distribution unlimited	
Supplementary Notes	
Abstract	
Subject Terms	
Document Classification unclassified	**Classification of SF298** unclassified
Classification of Abstract unclassified	**Limitation of Abstract** unlimited
Number of Pages 11	

coverage. Current operating frequencies for SAR sensors are limited in **their detection capability of targets embedded in deep foliage clutter. As a result a significant effort has been devoted to the development and demonstration of Foliage Penetration SAR sensors with the corresponding processing techniques.**

Multi and hyper spectral sensors provide complementary information by measuring solar reflection and the thermal emission across targets and backgrounds. The result is multiple 2D projection of the object and its internal attributes (as a function of wavebands) onto a plane perpendicular to the line of sight. These sensors can provide high definition imagery of the object silhouette, provided that there is discernible contrast (in an individual band or accumulated over bands) between target and background. High resolution spectral sensors can also provide definition of individual target features such as gun, hatch, wheels, engine compartments, exhaust ports, etc., based on color, intensity or thermal differences between target components**.** We have seen sensors operating at the reflective region of the EO spectrum from visible wavelengths to short-wave IR (SWIR), as well as thermal sensors operating at either midwave (3-5µ) or longwave (8–12µ). Multi-spectral (MS) or hyper-spectral (HS) sensors cover the EO to far IR spectral region typically with 10 to 15 broad bands for MS sensors and 50 to hundreds of fine, narrow bands for HS sensors. **In the reflective (visible to near IR) and the thermal (8 to 12µ) spectral region the added information can improve performance primarily against camouflage, concealment, and countermeasures.**

Each of the two sensors can provide robust ATC performance and defeat CC&D, but the false alarm rate is higher than desired. The fusion of information extracted from these two sensors will provide the desired ATC/ATR performance, while minimizing the false alarm rate to a manageable level. This is due to the complementary information across the two sensing domain and the differences in the source of false alarms.

2.0 FUSION METHODS

Four levels of multiple sensor data fusion are described. The highest level of fusion occurs when the multiple images are combined to a single image. Each location in the combined image has an associated vector of measurements from each of the sensors. The new image is then processed by an algorithm (such as an ATC/R) that simultaneously operates on the vector of values. Fusion techniques that operate in that mode are known as **centralized data fusion methods**. They typically assume common image projection plane for the multiple sensors and often rely on high level of image correlation. The centralized technique contains a fusion center in which the measurements or feature vectors from each of the sensors are processed to form a global decision. One such sensor configuration that is well matched for this type of approach is a multi-spectral or hyper-spectral sensor. In this case, the data is highly correlated for natural clutter and well aligned, thus providing the capability for **fusing** the image data **at the pixel-level**.

The next level of sensor fusion is **distributed data fusion**. In distributed fusion, each sensor makes an independent decision based on its own observations and passes these decisions to the fusion node where a global decision is made. Since the distributed fusion technique transmits less information to the fusion node, its performance may be degraded relative to the centralized approach. This approach provides a more practical solution to near real-time systems, and it offers maximum benefit when there is simultaneous sensing by the various sensors.

An alternate approach that utilizes the advantages from both the centralized and distributed techniques is termed **hybrid fusion**. Figure 1 illustrates the hybrid fusion architecture. First, each sensor makes an independent report based on its own observations or features. The process of automatic target cueing may consist of only the first stage of detection or may incorporate the second stage of target discrimination (false alarm reduction) for each of the sensors before the individual decisions are made. Thus, a list of candidate targets is independently generated from each sensor. This preliminary detection hypothesis is a soft decision. The combined hypothesis space is focused only on candidate targets that appear in both image domains. This results in a significant reduction in the number of hypotheses for subsequent processes. It also simplifies the geolocation/registration process, since there is only a need to associate cues.

This approach is applicable only if the probability of detection is high in each sensor domain, otherwise the probability of detection will be driven by the lowest performing sensor. If the false alarms are not correlated in terms of location, a substantial reduction will be achieved based only on geolocation. However, even if the false alarms correlate in terms of position, a substantial reduction is still achievable if the extracted features decorrelate. In this case the joint distribution of the extracted features will provide better separability (target/clutter) than each feature set alone.

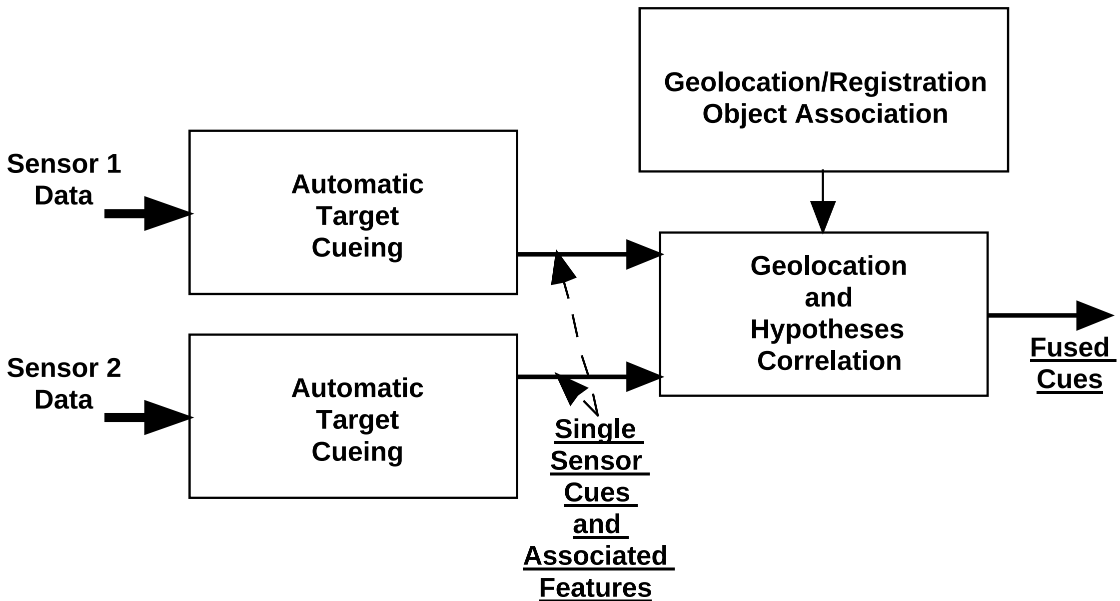


Figure 1: Hybrid Data Fusion Architecture

The next level of fusion, termed **cued fusion** (depicted in Figure 2), designates one sensor as the primary sensor and utilizes the second sensor for false alarm reduction. In this mode of fusion, the data of the primary sensor is used to derive an initial set of target cues, a set of associated features/attributes and the level of confidence based on the primary sensor data. Since ATC is performed only in the primary sensor domain, the corresponding image location for each target cue in the other sensor domain is estimated by automated geolocation/registration processes. A set of target features is derived from the estimated target location and combined with the attributes extracted from the primary sensor. The combined set of features is passed to a classifier that is trained on the joint distribution. A final decision is made by combining the estimated target confidence derived from the primary sensor and the confidence based on the joint set of features. The cued fusion requires either better geolocation accuracy than the hybrid approach or an alternate mechanism for refining geolocation.

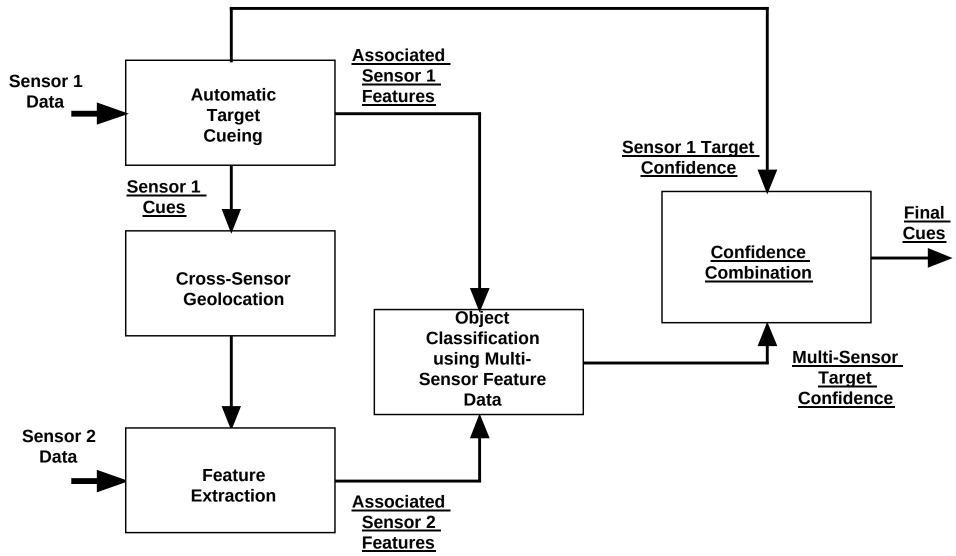


Figure 2: Cued Fusion Architecture

The lowest/minimal level of sensor fusion, termed **hand-off fusion**, is a mode in which one sensor cues a second sensor to an area of interest. At this point a complete hand-off to the secondary sensor occurs, and all of the imagery collected by this sensor is independently processed. The final set of detection cues is generated from data of the cued sensor.

3.0 FUSION METHODS FOR SPECTRAL AND SAR SENSORS

Two of the five methods that were described (**hybrid fusion and cued fusion**) present a practical fusion process that provides the performance improvements for SAR and hyper/multi spectral (HS/MS) sensors. For SAR and HS sensors, simultaneous sensing over the same region on the ground is likely to occur when the two sensors are mounted on separate platforms. If the two sensors are located on a single platform, the two sensors may image the same region with a time delay. SAR sensors are operating at standoff ranges (side looking), while HS sensors can image from nadir to lower depression angles. Since the sensing geometries are likely to be different, the fusion node will have to correlate the decisions of each sensor. In this case, if the targets are over resolved (even at coarse to medium resolutions), the correlation does not have to achieve the accuracy required by the centralized fusion process.

The backbone of the hybrid and cued fusion processes is a target/clutter classifier that either operates against features extracted from a single sensor or against the combined multi-sensor feature set. Figure 3 illustrates the overall target/clutter discrimination process that is applied to the set of features. It consists of two main processes: feature extraction and detection classification. The selected set of features is aimed at achieving discrimination between targets and natural clutter. A training set is used to estimate target/clutter statistics required for actual operation, and the ground truth information is used to determine whether the object is a target or a clutter for the training portion of the algorithm. For performance evaluation and actual operation the extracted features generated from either test imagery (not included in the training set) or actual imagery are processed by the classifier to determine the class (target, clutter) of the detected objects and the confidence level for that decision.

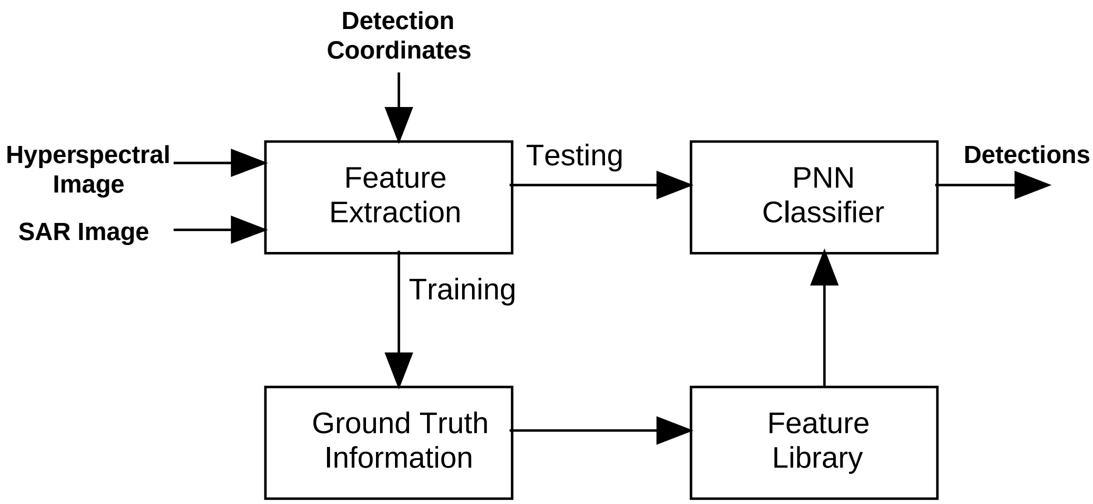


Figure 3: Feature-level Fusion Block Diagram

3.1 Cross Sensor Correlation and Geolocation Accuracy

In order to perform feature level fusion (either hybrid or cued) there is a need to determine the corresponding location of each point/object on the ground in each of the input images. Because of inaccuracies in the geolocation process, an error basket proportional to the expected error has to be placed. More specifically, if a detection cue is located in one image at location (x1, y1), it is transformed to the estimated location (x2, y2) in the second image. These coordinates may not be physically located in the same region on the ground due to geolocation error. Therefore, an error basket is used to improve the likelihood that there is an object located within this uncertainty region in the second image.

3.2 Hybrid Fusion

There are three cases to be considered in hybrid fusion,

Case 1: No detection is generated in the combined hypothesis space when a detection is reported in one image but no detection is reported within the uncertainty region in the other image.

Case 2: If a single detection is reported within the uncertainty region of the other image, the features from both sensors are combined to the determine the object classification.

Case 3: If multiple detections are reported within the uncertainty region, the ambiguities are resolved through data association). We determine the one detection report that best matches the report from the other sensor. More specifically, we use the object discrimination/classification process to determine the correct data association. For example, if for a single detection in the spectral image there are three detections reported in the SAR uncertainty region. A feature vector is generated for all of the three SAR objects. Another set of features is generated for the single object in the spectral image domain. These features are combined with each of the three SAR feature vectors to form three distinct combined feature vectors. Each of these feature vectors is examined by the pre-trained target/clutter classifier to determine if any of them exhibits target-like characteristics. If all of them are classified as clutter, the detection cue generated at this location is rejected as a false alarm. Otherwise, a detection is retained with the corresponding combined set of features that exhibited the strongest target-like behavior.

3.3 Cued Fusion

In cued-level fusion, we can not use the object discrimination approach to correct for geolocation errors, since it performs ATC only on the primary sensor. Normally, this approach requires very accurate geolocation in order for the feature-level fusion process to perform well. If the transformed coordinates of the object in the primary image have a large geolocation error, the extracted features from the derived coordinates may not correspond to the detected object in the primary image. The larger the registration error, the worse the feature representation. However, it is possible to refine the initial estimate (within the error basket) using contrast. This approach has been evaluated under the MSET program and demonstrated that no significant degradation occurred with reasonable registration errors.

3.4 Feature Extraction

The key to the success of the two data fusion approaches is the feature-level fusion process as described in the previous section. The most important part of the feature-level fusion process is the selection of features for both spectral and SAR imagery that are used for target/clutter discrimination and to mitigate the effects of registration error.

The features can be grouped into three main categories: statistics-based, fractal-based and correlation-based. The statistics-based features generally use amplitude-based statistics to characterize the detected area. Fractal-based features estimate the fractal/no-fractal behavior. The correlation-based features measure the level of spatial

correlation of targets and clutter. These features are combined into a single feature vector that is passed to the pre-trained classifier to eliminate/reduce false alarms.

3.5 Feature Classification

A classifier is used to determine the separability between targets and clutter. One common method is to apply a minimum distance classifier that is equivalent to the Bayesian linear classifier when the feature vector elements are uncorrelated, Gaussian, and have unit variance. Assuming that both targets and clutter are equally likely, the decision is target if

$$P_1(\mathbf{F}) < P_2(\mathbf{F}) \tag{1}$$

where

$$P_i(\mathbf{F}) = (\mathbf{F} - \overline{\mathbf{m}}_i)(\mathbf{F} - \overline{\mathbf{m}}_i)^T$$

$\mathbf{m}_i$ = mean feature vector of the ith class ($\mathbf{m}_1$ - mean of target class, $\mathbf{m}_2$ - mean of clutter class)
$\mathbf{F}$ = feature vector to be classified

The assumption in Eq. 1 ignores any correlation across features, however the band-to-band correlation exists in hyperspectral image data as it is exploited to accomplish improved clutter suppression. To account for the band-to-band correlation the general form of the Bayesian linear classifier, assuming normal distributions, is given by

$$P_i(\mathbf{F}) = -0.5(\mathbf{F} - \overline{\mathbf{m}}_i)\overline{\Sigma}_i^{-1}(\mathbf{F} - \overline{\mathbf{m}}_i)^T - 0.5\ln\left|\overline{\Sigma}_i\right| \tag{2}$$

where $\mathbf{m}_i$ = mean feature vector of the ith class
($\mathbf{m}_1$ - mean of target class, $\mathbf{m}_2$ - mean of clutter class)
Σ_i = covariance matrix of each class
$|\Sigma|$ = determinant of the covariance matrix

$\mathbf{F}$ = feature vector to be classified

For the combined spectral/SAR feature vector, Σ_i becomes,

$$\Sigma i = \left| \begin{matrix} \Sigma i_{HS} & 0 \\ 0 & \Sigma i_{SAR} \end{matrix} \right| \tag{3}$$

3.6 PNN Classifier

Equation (2) assumes that the features have a Gaussian distribution. This assumption is not generally valid for real clutter. In addition, for a large number of features, such as might be the case for the combination of spectral and SAR sensor data, the inverse of the covariance matrix may not have a stable solution (matrix of lower rank than the number of features). These two characteristics of classical classifiers have motivated the selection of a different type of classifier to perform multi-sensor target/clutter discrimination. More specifically, a modified version of a Probabilistic Neural Network (PNN) classifier was selected. It can be shown that it maps into a feed-forward neural network structure typified by many simple parallel processors. Unlike a variety of neural network classifiers the training time for the PNN is negligible. The PNN classifier can be expressed as a sum of Gaussian probability density functions as shown below:

$$\mathrm{P}_i(\mathbf{F}) = \frac{\mathrm{K}}{\mathrm{m}_i} \sum_{i=1}^{m_i} \exp[-\frac{1}{2}(\mathbf{F} - \mathbf{X}_{ij})^{\mathrm{T}} \left(\alpha \Sigma_i\right)^{-1} (\mathbf{F} - \mathbf{X}_{ij})] \tag{4}$$

where $K = (2\pi)^{-n/2}|\Sigma_i|^{-1/2}$

i = class number

j = training pattern number

m_i = total number of training patterns

$\mathbf{X}_{ij}$ = jth training pattern from class i

n = number of features in the pattern (dimension space)

Σ_i = covariance matrix of class i

α = covariance gain factor

$\mathbf{F}$ = feature vector to be classified

$P_i(\mathbf{F})$ is simply the sum of small multivariate Gaussian distribution centered on each training sample. However, the sum is not limited to being Gaussian. In fact it can approximate any smooth density function. The decision boundary of the probability density function has been shown to asymptotically approach the Bayesian optimal decision surface as the number of training patterns increases. For a large feature set, the inverse covariance may also become unstable, thereby degrading performance. The standard Quadratic Bayesian classifier exhibits this effect too. In order to avoid the inversion of the covariance matrix, only the variance of the individual feature distribution (with a user defined scaling factor) is used. Figure 4 provides an example of hybrid feature level fusion. It shows the detection cues generated by AAEC's geometric whitening filter (GWF) and overlaid on one of the multi-spectral bands and SAR-derived detection cues overlaid on the SAR image. The combined fusion process was able to detect the two camouflaged targets with a single false alarm.

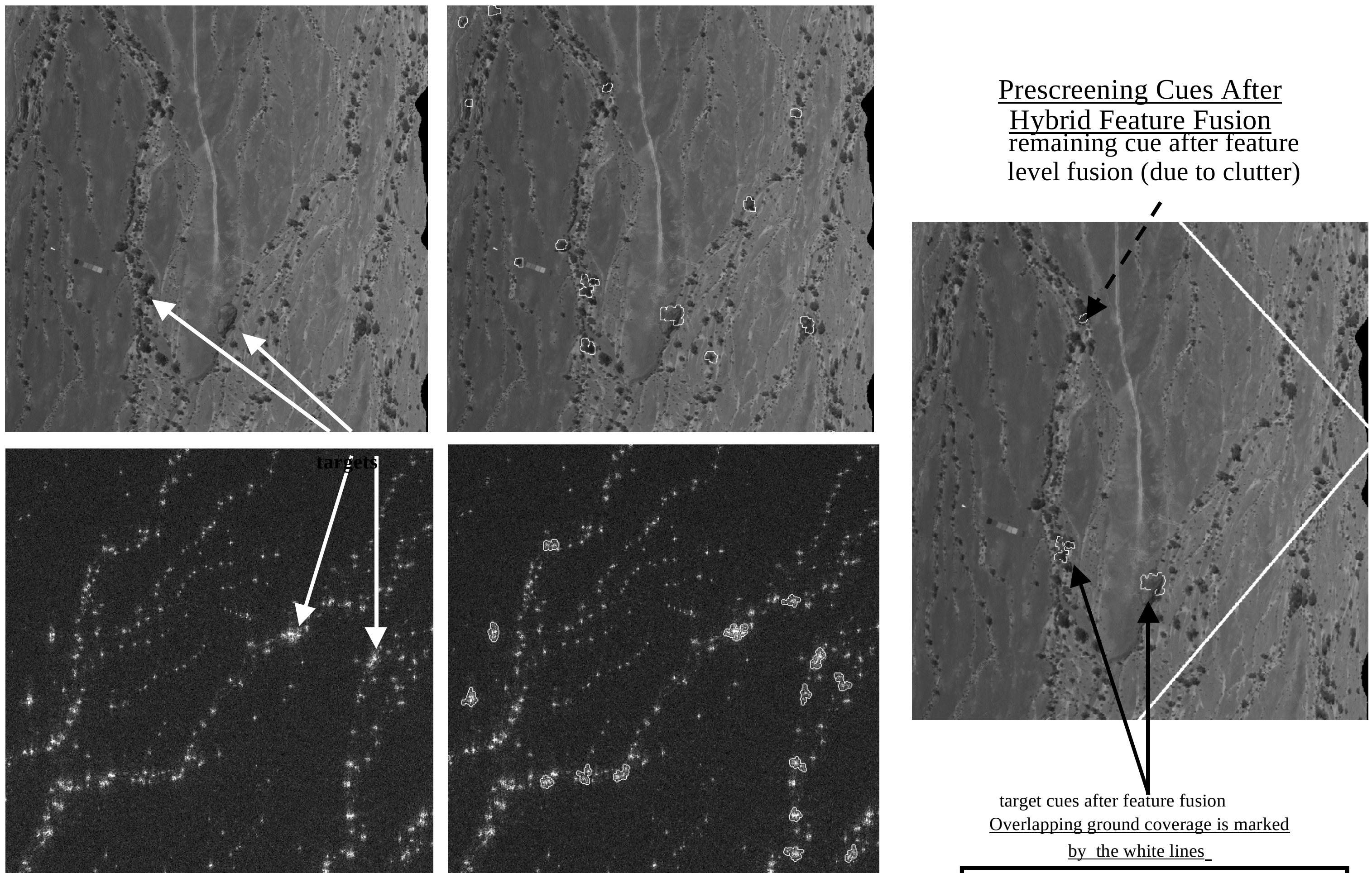
ACKNOWLEDGEMENTS

This work was performed under the sponsorship of the DARPA under SBIR Contract No. DAAH01-97-C-R151. We wish to thank Dr. Mark Davis and Lt. Col. Brad Tousley for their support.

REFERENCES

1. Peli, T., "Multiscale Fractal Theory and Object Characterization", Optical Society of America A, Vol. 7, No. 6, June, 1990.
2. A. Papadopoulos, M. Young, *A synergistic Centralized Multi-Sensor Fusion Technique For Ground Mobile Target Detection*, June 20, 1990.
3. Peli, T.,Young, M., and Ellis, K" Combining Linear and Nonlinear Processes for Multi-Spectral Material Detection/Identification," SPIE Conf. on Algorithms for Multispectral and Hyperspectral Imagery II, 1997
4. Stotts, L.B., Winter, E.M. and Reed, I.S., "Clutter Rejection Using Multispectral Processing," SPIE Aerospace Sensing, Conference 1305 (16-20 April) 1990, Orlando, FL.

Figure 4 : Hybrid feature level fusion for false alarm reduction using
SAR and MS (5 bands) Imagery



target cues after feature fusion

Overlapping ground coverage is marked
by the white lines

Multiple sensor inputs are at
approximately the same scale, but
the sensing geometries are different

raw data : top - single spectral band (1 of 5) output of single sensor prescreening
bottom - SAR imagery (cues appear as white overlays)